



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/779,764	02/18/2004	Yun-Wen Lee	0941-0917P	1610
2292 7590 05/21/2009 BIRCH STEWART KOLASCH & BIRCH PO BOX 747 FALLS CHURCH, VA 22040-0747				
EXAMINER BORSETTL GREG				
ART UNIT 2626		PAPER NUMBER		
NOTIFICATION DATE 05/21/2009		DELIVERY MODE ELECTRONIC		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

mailroom@bskb.com

Office Action Summary

Application No.

10/779,764

Applicant(s)

LEE, YUN-WEN

Examiner

GREG A. BORSETTI

Art Unit

2626

Period for Reply -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 01 December 2008.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1, 2, 6, 7, 9, 10, 14-16 and 18-20 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-2, 6-7, 9-10, 15-16, 19-20 is/are rejected.
- 7) ☒ Claim(s) 14 and 18 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 12/10/2008 has been entered.

Response to Amendment

2. Claims 1, 2, 6-7, 9-10, 14-16, and 18-20 are pending.
3. Claims 1, 6, 9, 14-15 have been amended.
4. Claims 3-5, and 11-13 have been canceled.
5. Claims 18-20 have been added.

Response to Arguments

6. In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., fixing left/right contexts and mapping right/left contexts obtaining the multi-lingual mapping set, Page 9, ¶ 4) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read

into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

7. Applicant's further arguments with respect to the claims have been considered but are moot in view of the new ground(s) of rejection.

Claim Objections

8. Claim 1 is objected to because of the following informalities: Claim 1 recites "to generates a speech command" (last limitation). It should be rewritten to read "to generate a speech command". Appropriate correction is required.

9. Claims 2 and 10 objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim(s), or amend the claim(s) to place the claim(s) in proper dependent form, or rewrite the claim(s) in independent form. Claims 2 and 10 further limit the speech models to be diphone models. However, amended claims 1 and 9 now claim that the models are generated by a diphone model generation engine, meaning that they are diphone models. Appropriate correction is required.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

10. Claim(s) 9-10,14-16, and 20 is/are rejected under 35 USC 101 for being

nonstatutory. Under the most recent interpretation of the Interim Guidelines regarding 35 U.S.C. 101, a method claim must (1) be tied to another statutory class or (2) transform underlying subject matter to a different state or thing. If no transformation occurs, the claim(s) should positively recite the other statutory class to which it is tied to qualify as a statutory process under 35 U.S.C. 101. As for guidance to areas of statutory subject matter, see 35 U.S.C. 101 Interim Guidelines (with emphasis of the Clarification of "processes" under 35 USC 101); As an example, the claim(s) could identify the apparatus that accomplishes the method steps, or positively recite the subject matter that is being transformed.

As per independent claim 9, the claim may be interpreted as a human manually performing the method of recognizing a mixed multi-lingual speech signal; comparing a plurality of multi-lingual query commands to obtain a plurality of multi-lingual baseforms; selecting and combining the multi-lingual baseforms by fixing one side context and mapping the other to get a mapping results; obtaining the multi-lingual context-speech mapping data according to the mapping result; writing down the result; locating and comparing a plurality of candidate data sets corresponding to the speech features according to the multi-lingual model database to find match probability of a plurality of candidate speech models of the candidate data sets; and selecting a plurality of resulting speech models corresponding to the speech features from the candidate speech models according to the match probability to generate a speech command.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

11. Claims 1-2, 9-10, and 19-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over D'Hoore (US Patent #6085160) in view of Burns (US Patent #5454106) and further in view of Black et al. (NPL Document "Building Voice in the Festival Speech Synthesis System")

As per claim 1, D'Hoore teaches:

a speech modeling engine, receiving and transferring a mixed multi-lingual speech signal into a plurality of speech features (column 2 lines 7-13 and column 3 lines 34-40);

a multi-lingual baseform mapping engine, comparing a plurality of multi-lingual query commands to obtain a plurality of multi-lingual baseforms; (column 3 lines 9-18, the system recognizes phonemes or phoneme like units, therefore it is inherent that the system first performs tokenization, or obtains baseforms); and

a cross-lingual model generation engine, coupled to the multi-lingual baseform mapping engine, selecting and combining the multi-lingual baseforms, further

comprising: (column 3 lines 22-25 and column 4 line 63 – column 5 line 14, context dependent biphone acoustic models are trained and used for recognition).

a speech search engine, coupled to the speech modeling engine, receiving the speech features, and locating and comparing a plurality of candidate data sets corresponding to the speech features, according to the multi-lingual model database to find match probability of a plurality of candidate speech models of the candidate data sets (column 4 lines 42-45, feature vectors are compared to acoustic models (connecting sequences) and then to a multilingual acoustic model (multi-lingual model database) to determine the best match); and a

decision reaction engine, coupled to the speech search engine, selecting a plurality of resulting speech models corresponding to the speech features according to the match probability from the candidate speech models to generates a speech command (column 4 lines 42-45, feature vectors are compared to acoustic models and then to a language model to determine the best match).

D'Hoore fails to teach, but Burns teaches:

Burns discloses inputting query commands to a speech recognizer, which are then sent to be scanned by a tokenizer (column 4 lines 20-29). Burns discloses a system that enables a user to retrieve information from a database using natural language queries (column 3 lines 10-15).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to compare a plurality of multi-lingual query commands to obtain a

plurality of multi-lingual baseforms in D'hoore, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp in order to achieve the predictable result of producing a multi-lingual speech recognition system optimized for a variety of recognition tasks.

D'Hoore and Burns fail to teach but Black teaches:

diphones as sub-word units; (section 5)

fixing one side contexts of the multi-lingual baseforms and mapping another side contexts of the multi-lingual baseforms to obtain a mapping result; (Section 3.1.2, ...*Some pronunciations change depending on the context they are in...* Section 5.1-5., ...*Diphone synthesis and in general any concatenative synthesis method make an absolute fixed choice about which units exist and in circumstances where something else is required a mapping is necessary...* Specifically Page 26 mentions multi-lingual mappings. Page 45, teaches a left context where it would have been obvious to someone of ordinary skill in the art at the time of the invention that a left context would change the right mapping based on the context.)

obtaining the multi-lingual context-speech mapping data according to the mapping result; and (Sections 5.1 – 5.2 teaches contextual dependencies on how diphones are generated and used for multi-lingual pronunciations. Page 24 shows a generated list for the mapping which is used for synthesis.)

storing the multi-lingual context-speech mapping data in a multi-lingual model database; (Section 6.2, the diphone models are stored in a database)

It would have been obvious to someone of ordinary skill in the art at the time of the invention to combine Black with D'Hoore and Burns to consider foreign phone pronunciations because "in most languages nowadays making no attempt to accommodate foreign phones is considered ignorant at least and possibly even arrogant" (Page 23)

As per claims 2 and 10, claims 1 and 9 are incorporated and D'Hoore and Burns fail to teach but Black teaches:

wherein the speech models are characterized by diphone models.
(section 5)

It would have been obvious to someone of ordinary skill in the art at the time of the invention to combine Black with D'Hoore and Burns to consider foreign phone pronunciations because "in most languages nowadays making no attempt to accommodate foreign phones is considered ignorant at least and possibly even arrogant" (Page 23)

As per claim 9, D'hoore teaches:

transferring a mixed multi-lingual speech signal into a plurality of speech features; (column 2 lines 7-13 and column 3 lines 34-40);

comparing a plurality of multi-lingual query commands to obtain a plurality of multi-lingual baseforms; (column 3 lines 9-18, the system recognizes phonemes or phoneme like units, therefore it is inherent that the system first performs

tokenization, or obtains baseforms);

selecting and combining the multi-lingual baseforms, comprising: (column 3 lines 22-25 and column 4 line 63 – column 5 line 14, context dependent biphone acoustic models are trained and used for recognition).

locating and comparing a plurality of candidate data sets corresponding to the speech features according to the multi-lingual model database to find match probability of a plurality of candidate speech models of the candidate data sets; and (column 4 lines 42-45, feature vectors are compared to acoustic models (connecting sequences) and then to multilingual acoustic model (multilingual model database) to determine the best match)

selecting a plurality of resulting speech models corresponding to the speech features from the candidate speech models according to the match probability to generate a speech command. (column 4 lines 42-45, feature vectors are compared to acoustic models and then to a language model to determine the best match).

D'Hoore fails to teach, but Burns teaches:

Burns discloses inputting query commands to a speech recognizer, which are then sent to be scanned by a tokenizer (column 4 lines 20-29). Burns discloses a system that enables a user to retrieve information from a database using natural language queries (column 3 lines 10-15).

Therefore it would have been obvious to one of ordinary skill in the art at the time

of the invention to compare a plurality of multi-lingual query commands to obtain a plurality of multi-lingual baseforms in D'hoore, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp in order to achieve the predictable result of producing a multi-lingual speech recognition system optimized for a variety of recognition tasks.

D'Hoore and Burns fail to teach but Black teaches:

fixing one side contexts of the multi-lingual baseforms and mapping another side contexts of the multi-lingual baseforms to obtain a mapping result; (Section 3.1.2, *...Some pronunciations change depending on the context they are in...* Section 5.1-5., *...Diphone synthesis and in general any concatenative synthesis method make an absolute fixed choice about which units exist and in circumstances where something else is required a mapping is necessary...* Specifically Page 26 mentions multi-lingual mappings. Page 45, teaches a left context where it would have been obvious to someone of ordinary skill in the art at the time of the invention that a left context would change the right mapping based on the context.)

obtaining the multi-lingual context-speech mapping data according to the mapping result; and (Sections 5.1 – 5.2 teaches contextual dependencies on how diphones are generated and used for multi-lingual pronunciations. Page 24 shows a generated list for the mapping which is used for synthesis.)

storing the multi-lingual context-speech mapping data in a multi-lingual model database; (Section 6.2, the diphone models are stored in a database)

It would have been obvious to someone of ordinary skill in the art at the time of the invention to combine Black with D'Hoore and Burns to consider foreign phone pronunciations because "in most languages nowadays making no attempt to accommodate foreign phones is considered ignorant at least and possibly even arrogant" (Page 23)

As per claims 19 and 20, claim 1 and 9 are incorporated and D'hoore teaches:

wherein the speech search engine locates and compares the candidate data sets, further referring the connecting sequences of the speech features and a speech rule database. (column 4 lines 42-45, feature vectors are compared to acoustic models (connecting sequences) and then to a language model (speech rule database) to determine the best match);

12. Claims 6,7,15 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over D'Hoore (US Patent #6085160) in view of Burns (US Patent #5454106) and further in view of Black et al. (NPL Document "Building Voice in the Festival Speech Synthesis System") and further in view of Waibel ("Interactive Translation of Conversational Speech" IEEE 1996).

As per claims 6 and 15, D'hoore, Burns, and Black disclose the system as claimed in claims 1 and 9, however D'hoore does not disclose wherein the multi-lingual model database comprises a plurality of multi-lingual anti-models. Waibel discloses a system

for speech recognition which uses garbage models to model nonstationary noises (page 44, second paragraph). These garbage models, also known as anti-models, are used to model common nonspeech noises, such as coughs and lip-smacking, and non human noises, such as a door slams and ringing telephones.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use anti-models in D'hoore, Burns, and Black, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp in order to achieve the predictable result of removing non-speech and background noises, thus improving the overall recognition accuracy.

As per claims 6 and 15, D'hoore, Burns, and Black disclose the system as claimed in claims 1 and 9, but D'hoore does not disclose at least one uni-lingual anti-model generation engine, receiving a plurality of multi-lingual query commands to generate a plurality of uni-lingual anti-models corresponding to specific languages; and an anti-model combination engine, coupled to the uni-lingual anti-model generation engine, calculating the uni-lingual anti-models to generate the multi-lingual anti-models. However, D'hoore does disclose receiving multi-lingual speech input and training multi-lingual acoustic models (column 4 line 63- column 5 line 14). In addition, Waibel discloses a system for speech recognition which uses garbage models to model non-stationary noises (page 44, second paragraph). These garbage models, also known as anti-models, are used to model common non-speech noises, such as coughs and lip-smacking, and non human noises, such as a door slams and ringing telephones.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to receiving a plurality of multi-lingual query commands to generate a plurality of uni-lingual anti-models corresponding to specific languages, and use an anti-model combination engine, coupled to the uni-lingual anti-model generation engine, to calculate the uni-lingual anti-models to generate the multi-lingual anti-models in D'hoore, Burns, and Black, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp in order to achieve the predictable result of removing non-speech and background noises, thus improving the overall recognition accuracy.

Allowable Subject Matter

13. Claim 14 and 18 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.
14. The following is a statement of reasons for the indication of allowable subject matter:

As per claim 14, claim 9 is incorporated and the closest known prior art fails to teach alone or in fair combination:

fixing left contexts of the multi-lingual baseforms and mapping right contexts of the multi-lingual baseforms to obtain a mapping result;

fixing right context and mapping the left contexts of the multi-lingual baseforms to

obtain the mapping result if the right contexts of the multi-lingual baseforms mapping fails; and

obtaining the multi-lingual context-speech mapping data according to the mapping result.

D'Hoore teaches a multi-lingual speech recognition system but fails to teach multi-lingual mapping based on context. Burns teaches query commands to a speech recognition system. Lastly, Black teaches contextual mapping but fails to teach that the left context is fixed then right mapping and if that fails then fixing the right context and mapping the left to recognize multi-lingual speech.

Claim 14 would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims and also overcoming the 35 USC 101 rejection in claim 9.

As per claim 18, claim 1 is incorporated and the closest known prior art fails to teach alone or in fair combination:

fixing left contexts of the multi-lingual baseforms and mapping right contexts of the multi-lingual baseforms to obtain a mapping result;

fixing right context and mapping the left contexts of the multi-lingual baseforms to obtain the mapping result if the right contexts of the multi-lingual baseforms mapping fails; and

obtaining the multi-lingual context-speech mapping data according to the

mapping result.

D'Hoore teaches a multi-lingual speech recognition system but fails to teach multi-lingual mapping based on context. Burns teaches query commands to a speech recognition system. Lastly, Black teaches contextual mapping but fails to teach that the left context is fixed then right mapping and if that fails then fixing the right context and mapping the left to recognize multi-lingual speech.

Claim 18 would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims and also overcoming the objection in claim 1.

Conclusion

15. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Refer to PTO-892, Notice of References Cited for a listing of analogous art.

16. Any inquiry concerning this communication or earlier communications from the examiner should be directed to GREG A. BORSETTI whose telephone number is (571)270-3885. The examiner can normally be reached on Monday - Thursday (8am - 5pm Eastern Time).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, RICHEMOND DORVIL can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-

Art Unit: 2626

273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Greg A. Borsetti/
Examiner, Art Unit 2626

/Talivaldis Ivars Smits/
Primary Examiner, Art Unit 2626

5/18/2009